

3分木による繰り返しじゃんけん戦略の表現と進化的獲得

メタデータ	言語: Japanese 出版者: 公開日: 2011-04-05 キーワード (Ja): キーワード (En): 作成者: 牧野, 泰裕, 小高, 知宏, 小倉, 久和, Yasuhiro, MAKINO, Tomohiro, ODAKA, Hisakazu, OGURA メールアドレス: 所属:
URL	http://hdl.handle.net/10098/3115

3分木による繰り返しじゃんけん戦略の表現と進化的獲得

牧野 泰裕* 小高 知宏** 小倉 久和**

Representation and Evolutional Acquisition of Repeated Janken Game Strategy by 3-Branch-Tree Structure

Yasuhiro MAKINO*, Tomohiro ODAKA** and Hisakazu OGURA**

(Received February 19, 2003)

In this paper, we investigate the repeated janken game (RJG) and examine the knowledge expression method. We acquire the strategy knowledge using genetic algorithm(GA). We propose 3-branch-tree as a representation of the knowledge structure and it is used to simulate RJG on computer. With these examinations we can see that it is effective method and will show usefulness of 3-branch-tree structure.

Key Words : RJG, GA, 3-Branch-Tree Structure

1. はじめに

本研究では、人がゲームの流れや状況においてどのように行動するか、その人の持つ交渉（思考）モデルを考える。ここでは、じゃんけんを繰り返して行なう繰り返し対戦型じゃんけんゲーム（Repeated Janken Game, 以下：RJG）^{[1]~[3]}を対象とし、戦略知識の獲得を試みる。

世界各地には似たような文化が存在し、じゃんけんと同じルールで RoShamBo^[4]というゲームの国際大会が行なわれている。この大会では、戦略をコンピュータのプログラムとして用意し、コンピュータ上で対戦する。戦略には、相手の手を利用したものや確率によるもの等、様々な戦略が用いられている。人は、じゃんけんという単純な交渉において、サイコロや乱数発生機などの機械的な装置を用いていなければ、何らかの方法、くせや好みを含めた意志決定法によって次の手を決めるため、手の出し方

に特徴がでるはずである。相手がもしある戦略モデルに従って次に出す手を決めているのならば、過去の手の系列から何らかの戦略モデルが推測できる可能性がある。

本研究では、戦略モデルの表現方法として、3分木構造を用いる。人間の記憶力は、過去5手分程度のパターンを表現できると考える。3分木構造を用いるのは、そのパターンを直観的に特徴を解りやすく捉えることができると考えたからである。モデルの獲得には遺伝的アルゴリズム（以下、GA）を用いる。

本論文では、3分木表現を用いたGAプレーヤと、人間の代用のダムプレーヤとで対戦を行った。実験では、GAプレーヤにより、ダムプレーヤの戦略に対して、最終的にその相手に勝ち越す戦略モデルを表現し、その特徴を捉えることを目指した。RJGにおける、3分木表現の検討のために実験を行った。その実験により得られた結果と考察について述べる。

2. RJGとその戦略知識獲得

本章では、RJGの対戦方法について説明する。さらに、GAを用いた戦略知識の獲得方法について述べる。

* 工学研究科システム設計工学専攻

** 知能システム工学科

* System Design Engineering Course, Graduate School of Engineering

** Dept. of Human and Artificial Intelligent Systems

2. 1 R J Gの設定

じゃんけんは、よく知られているように、グー(g), チョキ(t), パー(p)の三つの手を出し合って勝ち負けが決まるゲームである。1試合は、N回のじゃんけん対戦を1対1で行なう。(引き分けの場合も1回とカウントする。)1回の対戦で与えられる得点は、対戦結果により、勝ち+1点、引き分け0点、負け-1点である。2人のプレイヤーの手を、my-hand, opp-handとし、k回目 ($1 \leq k \leq N$)の対戦結果として出した手を、それぞれ $h[k]$, $o[k]$ と表す。 $h[k]$, $o[k]$, $k=1 \sim N$ は、ゲームの系列である。

2. 2 3分木による戦略表現

R J Gにおける戦略を獲得するためには、その知識を何らかの形で表現する必要がある。この節では表現方法である3分木構造表現について述べる。

3分木は、1つのノード(O)に、左からグー(G), チョキ(T), パー(P)の3つの枝(branch)を持つ。それぞれの枝の先には、葉(□)であるg, t, p, またはノードがその枝に結合する。これを1つの3分木の単位とし階層的に結合する。各層のノードには、my-handとopp-handの直前の過去4手前分の系列, $h_1 \sim h_4$, $o_1 \sim o_4$, のいずれかを割り当てる。結合深さは、人間の記憶力を考慮して7段までに制限した。例えば図1は次のような戦略モデルを示している。 o_1 ノードにより、opp-handの1手前の手がgならば(G)の先にある o_2 ノードの値がgの場合にpを出す。同様に、tの場合(T)の先の h_1 ノードの値がtの場合にgをだす。pの場合は、(P)の先の h_2 ノードの値がpの場合にtを出す。GAプレイヤーは、そのモデルにより決定された手を相手の手の推定手として、それに勝つ手を出す戦略をもつ。

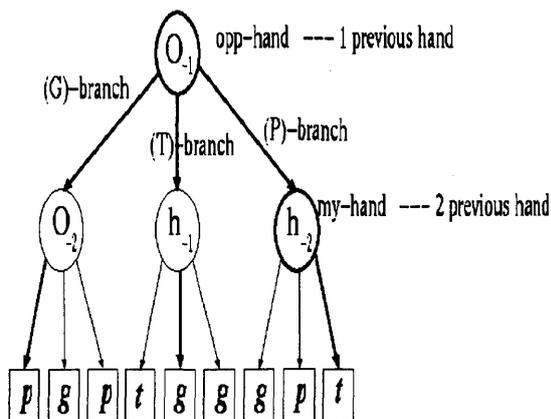


図1：3分木による戦略表現の例

2. 3 GAによる戦略知識獲得

GAによる相手の戦略知識獲得の流れを説明する。GAプレイヤーと対戦する相手のダムプレイヤーを用意する。この相手の戦略を「獲得対象戦略」と呼ぶ。この獲得対象戦略と対戦を繰り返すなかで、そのモデルをGAにより形成することを目的として、以下の処理を行なう。

(1) 1つの個体は1つの3分木構造で表される相手戦略のモデルである。GAプレイヤーは1つの個体の3分木表現されたモデルによる推定手に勝つ手を出す戦略を持つ。遺伝子プールとはいくつかの個体からなる進化的集団である。遺伝子プールの初期集団をランダムに生成する。

(2) 獲得対象戦略と遺伝子プールの中の各個体とで1試合ずつR J Gを行なう。各個体の評価値は1試合で得られたゲームの得点である。

(3) 打ち切り世代まで(4), (5)を繰り返し行なう。

(4) 遺伝子プール内の全ての個体について対戦を終えた後、遺伝子プールに対して交叉、突然変異を行なう。

(5) エリート保存戦略を併用してルーレット戦略で世代交代を行ない、個体集団を進化させ、獲得対象戦略を3分木表現モデルで形成する。

遺伝的操作は、交叉と突然変異を行なう。交叉は、まず2つの個体(3分木構造)を選択し、次にその部分木をランダムに選び、両者の部分木を取り換えて交叉を実行する。突然変異は、1つの個体を選択し、その部分木以下を新たに生成した部分木で置き換えたり、あるいは葉で置き換えることで実現する。

2. 3 ダムプレイヤーの戦略

ここでは、コンピュータシミュレーションのために、人間を想定したダムプレイヤーを用意した。ダムプレイヤーの獲得対象戦略として、考えられる戦略の中から、関数式($FS_1 \sim FS_3$)により決定する固定戦略と、3分木表現による戦略、対戦中の相手の手の履歴蓄積による履歴蓄積戦略のダムプレイヤーについて実験をした結果を示す。

2. 4 ダムプレーヤの戦略

(1) 固定的戦略

獲得対象とするダムプレーヤの戦略を次のような関数を用いて表現した。

my(x) : x 手前の my-hand の値を返す。
opp(x) : x 手前の opp-hand の値を返す。

次の手を決める関数式を FS とする。

式 FS では、常に計算した結果の3の剰余を用いており、0, 1, 2 は、それぞれ g, t, p に対応する。また、第1手目より前の手はランダムに決めた。次の戦略を用意した。

$$FS_1 = \text{my}(1) \quad (1)$$

$$FS_2 = \text{opp}(1) \quad (2)$$

$$FS_3 = (\text{my}(2) \times 2) + (\text{my}(1) + 1) \quad (3)$$

(2) 3分木表現戦略

3分木表現戦略は、図2に示す3分木表現モデルを用いて手を決める戦略である。

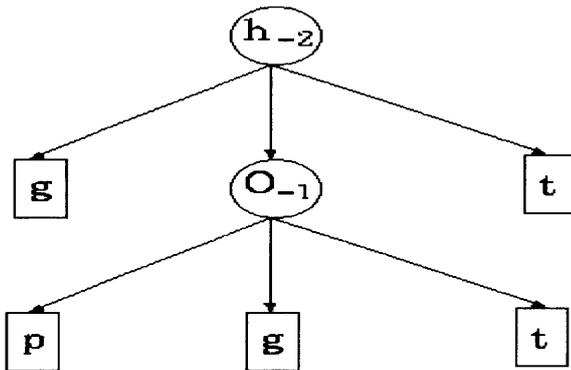


図2 : 3分木表現戦略

(3) 履歴蓄積戦略

履歴蓄積戦略は、関数式による戦略と違い、履歴の記憶の機能を持つ。R J Gの試合中のn回目の対戦で、履歴蓄積戦略のダムプレーヤ(my-hand)のn手h[n]を決める場合を考える。このとき、対戦相手であるGAプレーヤ(opp-hand)の履歴でk手分の手のパターン 3^k 通りのうち出現回数が一番多いパターンをGAプレーヤのn手目の推定手とする。履歴2蓄積戦略は、それに勝つようにする戦略である。今回は $k=3$ とした o_1, o_2 の2手分の履歴から次の手を推測する履歴2蓄積戦略を用いた。履歴2蓄積戦略は次のように定義できる。

まず、 $X, Y, Z \in \{g, t, p\}$ として要素 w_{XYZ} を、過去の履歴のうちにopp-handが手X, Yの次に手Zを出した回数とする。 w の初期値はすべて $w_{XYZ}=0$ とした。この w 'sにおいて一致するものを抜きだして、式(4)のように c_g, c_t, c_p とする。

$$\begin{cases} C_g = W_{o-2, o-1, g} \\ C_t = W_{o-2, o-1, t} \\ C_p = W_{o-2, o-1, p} \end{cases} \quad (4)$$

それぞれの出現回数 c_g, c_t, c_p の中で一番多い手に勝つ手を出す戦略が、履歴2蓄積戦略である。

3. 固定戦略のダムプレーヤとの実験結果

3. 1 実験の設定

3.1で説明したそれぞれの戦略を獲得対象戦略とし実験を行なった。実験においてGAプレーヤが得たのは、ダムプレーヤの戦略モデルであるから、oは、GAプレーヤ、hは、ダムプレーヤの手となる。その戦略モデルに勝つ手を出すのがGAプレーヤの手である。1つのダムプレーヤに対して、1試合(1000回)のR J Gを各個体(個体数=50)と行ない、打ちきり世代(2000)までGAによる獲得を行なった。実験は、乱数の初期値を変えて、交叉率=0.1, 0.2, 0.4, 0.6, 0.8と、突然変異率=0.001, 0.002, 0.004, 0.006, 0.008, 0.01, 0.02, 0.04, 0.06, 0.08の各設定で100回の実験を行なった。このなかで評価が高かった例を説明する。

3. 2 固定戦略ダムプレーヤとの戦略獲得実験

得られたモデルを用いた戦略と $FS_1 \sim FS_3$ の戦略に対して1試合のR J Gを行なった結果、それぞれに対し全勝する結果になった。太い枝はその試合時の推定に使われた部分である。以下に、そのモデルについて説明する。

図3に、 FS_1 に対して得られたモデルを示す。このモデルは、自分の手の1手前がgならばg, tならばt, pならばpを出し続けるモデルである。(T)の枝が今回使われた手であり、初期値がtの場合である。 FS_1 の戦略に対して得られた3分木表現のモデルは、my-handの1手前の手を用いる h_1 のノードからなる1段の3分木構造が得られた。初期値がg, t, pのそれぞれの場合に対して、 FS_1 の同じ手を出し続ける特徴を表現している。

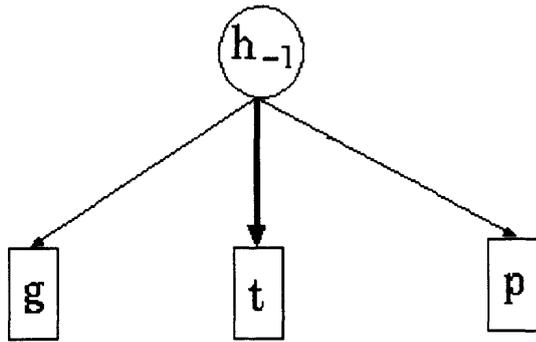


図 3 : FS₁ の 3 分木モデル

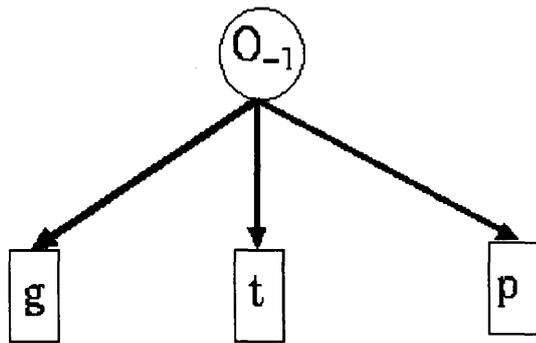


図 4 : FS₂ の 3 分木モデル

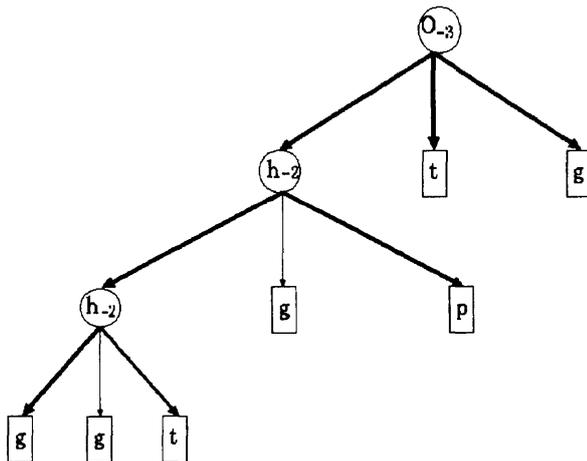


図 5 : FS₃ の 3 分木モデル

図 4 に、FS₂ に対して得られたモデルを示す。このモデルは、opp-hand の 1 手前の手を用いる o_{-1} のノードからなる 1 段の 3 分木表現である。FS₂ は相手の 1 手前と同じ手を仕返すというモデルである。

FS₃ の戦略に対しては得られた 3 分木表現モデルは、図 5 である。FS₃ の戦略は、1 つのパターンを繰り返し出す戦略であり、今回得られたモデルは、そのパターンを表現出来るモデルの 1 つとして、 h_{-2}, o_{-3} の 2 つの組合せの 3 分木構造により表現している。

関数式により決定する FS₁~FS₃(同じ手を出し続ける、相手の手を仕返す、パターンを繰り返す)戦略に対して、3 分木表現でモデル化して、その特徴を表現することができた。

3. 3 3 3 分木表現戦略プレーヤの戦略獲得実験

3 分木表現の戦略に対しては、図 5 と同じ戦略モデルが得られた。また、その他にも、 h_2 の初期値により簡略化したモデルも獲得できた。簡単な 3 分木表現の戦略に対しても、全く同じモデルと、初期値の場合によって簡略化されたモデルも獲得できた。

3. 4 履歴蓄積戦略のダムプレーヤとの実験

図 6 に、履歴 2 蓄積戦略に対して一番良い評価値を得られた 3 分木表現モデルを示す。履歴は 1 試合ごとに初期値 $w's=0$ としておこなった。このモデルを用いた戦略と履歴 2 蓄積戦略に対して 1 試合の R J G を行なうと、1000 回中 (741 勝 37 敗 222 分) と勝ち越せる。対戦履歴をみると、このモデルが出す手は、21 勝 1 敗 6 分となるようなパターンを繰り返していることが分かった。これは、履歴 2 蓄積の表現可能な 27 種類のパターンより長い長さの 28 のパターンを出すモデルである。このパターンは h_3, h_2, h_1 と、 o_3 を使って表現しているのがわかる。また、1 段目の h_3 が g の場合 g を推定する部分が、履歴 2 蓄積戦略の初期値の部分表現している部分である。

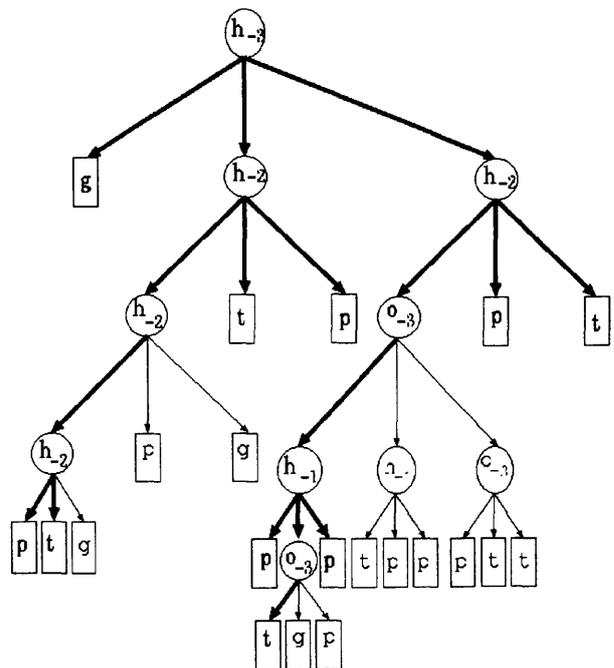


図 6 : 履歴 2 蓄積戦略の 3 分木モデル

履歴2蓄積戦略は、対戦相手の戦略に依存する戦略であるから、得られたモデルは履歴2蓄積戦略に対して、初期値の特徴を含んだ、負け越すように履歴蓄積させるモデルの一つである。

4. まとめ

本研究では、対戦型ゲームとしてR J Gを設定し、3分木表現を用いたGAプレーヤにより、人間の代用のダムプレーヤの戦略モデルの獲得を試みた。対戦相手には、関数式により決定する固定的戦略と、相手の蓄積履歴による履歴蓄積戦略を人間の代用としてのダムプレーヤとして用意した。結果から用意したダムプレーヤに対して勝ち越すことができた。これは、ダムプレーヤの戦略モデルを獲得できたからである。次の手を一意に決める固定的戦略 $FS_1 \sim FS_3$ に対して、3分木によりモデル化し特徴を表現できた。また、3分木表現による戦略のダムプレーヤに対しても、全く同じモデルだけでなく、簡略化した3分木モデルを表現できた。今回の例以外の表現についても対応が確認できたが、ここでは省略する。また、履歴蓄積戦略のように対戦相手の戦略に依存し変化する戦略に対しても実験をし結果、そのままのモデルは獲得できなかつた。それは、履歴蓄積戦略は相手の手により手の決定パターンが変化するために、そのすべてのパターンに対して実験するのは困難であるからである。しかし、履歴蓄積戦略に対しては、蓄積履歴の長さより長いパターンで勝ち越せるようなモデルを獲得できた。これにより、相互に影響する履歴戦略などにも対応できること確認できた。

過去の数手分の3分木という分かりやすい構造により、今回人間の代用としてのダムプレーヤに対して、相手の戦略モデルをGAで獲得できた。3分木表現の利点は、戦略モデルを一次元の記号列として解析し勝ち越すパターンのモデルを表現するだけでなく、木構造を用いている点である。木構造自体は、目新しいものではないが、履歴蓄積戦略に対して得られた結果のように、どの程度の履歴を考慮すれば勝ち越すようなパターンのモデルが得られるかといったことが確認できる。実際の人間の場合も、固定戦略のようではなく、履歴戦略のように相互作用の中で手が変化するとと思われる。よって、このような相手の戦略の特徴を表現するのに、3分木構造表現は有効であると考えられる。

現在、RJGにおいて、実際に人間の戦略モデルの獲得実験を試みている。実際の人間は、本論文で用意したダムプレーヤよりも複雑な戦略である。よっ

て、3分木表現の結合の深さや方法の改良をして、実際の人間に対しても効果を期待し、人間とGAプレーヤとでR J Gを行ない、対戦のなかで人間の戦略モデルの獲得を試みている

参考文献

- [1] 牧野泰裕, 西野順二, 小高知宏, 小倉久和: 対戦型ゲームにおける戦略知識獲得の試み, 平成8年度電気関係学会北陸支部連合大会講演論文集 p393, (1996).
- [2] 牧野泰裕, 西野順二, 小高知宏, 小倉久和: じゃんけん戦略の知識表現とその獲得, 平成12年度電気関係学会北陸支部連合大会講演論文集 p421, (2000).
- [3] 木村将治, 牧野泰裕, 小高知宏, 小倉久和: 繰り返しじゃんけんゲームを対象とした固定的戦略による知識表現と遺伝的アルゴリズムによる知識獲得, 電子情報通信学会論文誌, J85-D-I, No.5, pp.107--111, (2002.5). J82-D-II, No.3, pp.566--570, 1999.3
- [4] International Roshambo Programming Competition, <http://www.cs.ualberta.ca/~darse/rbspc.html>.
- [5] モートン・D・デービス: 「ゲームの理論」, 講談社 (1994)
- [6] 西山賢一: 「勝つためのゲーム理論」, 講談社 (1995).
- [7] 松原仁, 竹内郁雄: 「ゲームプログラミング」, 共立出版 (1997).
- [8] 飯田之弘之: 「プロ棋士のような戦略に基づくゲームプログラミングの試み」, 人工知能学会誌, Vol. 10, No. 6, p846, (1995).
- [9] 鈴木 光男: 「新ゲーム理論」, 勁草書房 (1994).
- [10] 北野宏明(編): 「遺伝的アルゴリズム」, 産業図書, (1993).

